

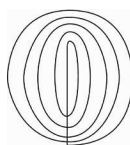
Versão PDF da entrada

IDEALIZAÇÃO DE MENTES

da EDIÇÃO DE 2014 do

COMPÊNDIO EM LINHA DE PROBLEMAS DE FILOSOFIA ANALÍTICA

2012-2015 FCT Project PTDC/FIL-FIL/121209/2010



Editado por
João Branquinho e Ricardo Santos

ISBN: 978-989-8553-22-5

Compêndio em Linha de Problemas de Filosofia Analítica
Copyright © 2014 do editor
Centro de Filosofia da Universidade de Lisboa
Alameda da Universidade, Campo Grande, 1600-214 Lisboa

Idealização de Mentes
Copyright © 2014 do autor
Helder Coelho

Todos os direitos reservados

Resumo

A criação de mentes, e o seu design, tem acompanhado a evolução do homem e a sua ânsia em replicar o poder da inteligência e de articular os seus principais componentes. Nos últimos 30 anos, a sua discussão (teoria da identidade, teorias da mente, governo versus controle) foi grande na Filosofia e, também, na Inteligência Artificial, sobretudo em redor dos agentes artificiais, em virtude da potencialidade de aplicações em áreas não muito afastadas, a saúde pública ou a sustentabilidade ambiental, a segurança e a energia, os fenómenos sociais ou as políticas. Nesta contribuição procurámos abrir um campo de reflexões e de conjecturas sobre a abstração da inteligência, introduzindo uma rede de conexões (memória) com termos (conceitos) aparentemente distantes, não esquecendo um olhar para os cérebros e os comportamentos. Simultaneamente, a imagem de mente ajudar-nos-á a invadir um território interdisciplinar e a confrontar/cruzar várias visões para simplificarmos ou complicarmos a ideia de mente possível. Recorrendo à Simulação Social, poderemos depois experimentar uma variedade de hipóteses, e deste modo percebermos que mesmo um agente artificial pode evoluir, graças a mentes mais complicadas, se quisermos ficar mais perto do realismo das situações.

Palavras-chave

Cognição, agentes inteligentes, arquiteturas, teorias da mente, simulação

Abstract

The creation of minds, and its design, has followed the evolution of man and the desire to replicate the power of his intelligence and to articulate his main components. Along the last 30 years, the discussion of minds (identity theory, theories of mind, govern versus control) was enlarged in Philosophy, and also in Artificial Intelligence, namely around artificial agents, and on account of its importance for supporting applications on healthcare, environment sustainability, security, energy, politics and policies. We tried to open the field of reflections and conjectures about the abstraction of intelligence by building up a network of terms, along the text, without forgetting brains and behaviors. The design of minds will help readers to invade the territory of other disciplines, crossing ideas, with the aid of social simulation and specific situations with less or more realism.

Keywords

Cognition, intelligent agents, architectures, theories of mind, simulation

Idealização de Mentes

1 Introdução

A ideia de mente está presente na Filosofia Analítica, na Inteligência Artificial (IA), na Ciência da Computação, nas Neurociências, na Ciência Cognitiva, ou mesmo na Biologia Animal, para identificar o que um cérebro faz, e associar o pensamento e o comando de um corpo (problema mente-corpo), graças aos seus sentidos (sensores) sobre o ambiente (paisagem de todos os possíveis objetos, onde os dados estão espalhados, e mergulhados em ruído e ambiguidade) que o cerca e, também, à consciência. Quando falamos em mentes, a ideia de teoria aparece a introduzir um modelo (Bratman 1987), uma arquitetura (Minsky 2006), uma estrutura (Sloman 1999), uma organização (Lenat 1995), um dispositivo, ou mesmo um mecanismo, caso do reconhecedor de padrões (Kurzweil 2012; Hawkins 2005). As mentes buscam os dados nas entradas e produzem cognições nas saídas (relações causais): Qual é o lugar para a Estatística, a estrutura e a abstração, perguntou Tenenbaum em 2011?

A mente contrapõe-se ao cérebro (veja-se a teoria da identidade), onde os estados e os processos estão interrelacionados, na tarefa de compreensão do mundo: o lado mental (representações, descrições, interpretações, experiências, comportamentos, sensações, percepções) e o lado físico (interações entre circuitos neuronais e biológicos).

Cérebro → Mente

Recentemente, no domínio da Evolução Humana, foram novamente comparadas as mentes dos Neandertal e dos Homo Sapiens no que respeita às suas capacidades cognitivas (e o eventual intervalo mental entre ambas), as diferenças do pensamento simbólico, a habilidade em aprender e memorizar as palavras, o recurso a ferramentas avançadas, a estratégia de procura de alimentos (caça de pequenos animais), ou o significado das diferenças anatómicas dos cérebros, incluindo as dimensões do lobo frontal, sede do poder de resolução de problemas. Concluiu-se, através de novas evidências arqueológicas, que as diferenças anteriormente encontradas poderão não ser tão

significativas como então se pensava, e que as razões da extinção dos primeiros não estão relacionadas com a inteligência.

Como se cria uma mente (Kurzweil 2012) para estar numa certa situação? E como se desenvolve o seu poder de agir? Putnam (1961) propôs vermos a mente como uma máquina de computação (processador de informação), e Fodor desenvolveu essas ideias (1975, 1980, 1987, 1993), destacando a razão e a causalidade, a necessidade de uma semântica dos estados mentais, e a cognição.

Mente \rightarrow Máquina (Mente Artificial)

Através de experiências de Simulação Social (Social Simulation e Social Computing) podemos, passo a passo, fazer crescer a inteligência de uma mente artificial através da sua melhor adequação às situações em que se coloca, desde a simples reação até à deliberação mais exigente. Ao modelamos os problemas (por exemplo, uma manifestação de rua, com ou sem violência) temos de preferir as arquiteturas (Lemos, Coelho e Lopes 2014) que sejam as mais apropriadas para as intervenções que queremos realizar (gritos ou murros, ou seja intervenção ou violência). Logo, a complicação só ocorre se for necessário aproximarmo-nos de um certo realismo. Daí, a profusão de teorias e de modelos, e alguma confusão quanto à sua necessidade e justificação, o que implica cuidado na verificação e validação dos modelos adotados.

A modelação baseada em agentes inteligentes e artificiais, que hoje em dia suporta a simulação social, é capaz de per si explicar o que se está a passar. Explica (faz crescer) regularidades macroscópicas sociais, as normas, os padrões espaciais, a dinâmica das infeções ou das instituições, a inteligência coletiva. A explicação geradora é também dedutiva (Epstein 2007). E, o desenho dos agentes permite a produção de comportamentos adequados, e é guiado por princípios, processos, dispositivos e mecanismos.

2 Modelos

Existem três tipos fundamentais de modelos (Gilbert 2007), abstrações ou descrições formais das components funcionais (o que), que são úteis para fins analíticos, que são correntemente considerados, quando se montam construções sociais, se geram animações ao longo

do tempo e se começa a fazer a experimentação recorrendo à modelação baseada em agentes (Agent Based Modeling ou ABM): os modelos abstratos, os de alcance médio e os de facsimile.

Ver: Estímulos \rightarrow Percepções

Ação: Percepção* \rightarrow Ações

Intenções \rightarrow Ações

Próximo estado: Intenção x Percepção \rightarrow Intenção

Nos primeiros visa-se demonstrar um processo (seletor) ou um mecanismo (planeador, decisor, filtro) social (provavelmente emergente), excluindo-se casos empíricos específicos, como na dinâmica de opiniões. O critério de validação respeita a dúvida se somos, nas experiências, conduzidos a teorias mais específicas que podem depois ser testadas empiricamente.

Nos segundos, o objetivo é descrever a característica geral de um fenómeno social particular, por exemplo, nos casos de inovação tecnológica, de bom comportamento (via regras morais) ou nas cadeias de fornecimento de produtos. O critério de validação é agora a semelhança qualitativa, a dinâmica, ou as situações históricas amigáveis.

Nos terceiros, pretende-se fornecer uma reprodução exata de um certo fenómeno, seja no caso do trânsito de uma cidade, onde o alvo é prever locais potenciais de engarrafamento. Agora, o critério da validação é o realismo que pretendemos atingir. Existe um problema do comportamento deste modelo pois ele pode ser fortemente influenciado por eventos aleatórios (a simulação pode recorrer a um gerador de números aleatórios).

Estes tipos de modelos colocam questões quanto à verificação (Será o modelo correto, mesmo após nos vermos livres dos erros que cometemos) e à validação (Será o modelo bom, o que está dependente dos objetivos, como 1) a formalização de uma teoria graças a um modelo abstrato, 2) o desenvolvimento de uma teoria de médio alcance, graças a um modelo de uma classe de fenómenos, ou a descrição de uma situação específica graças a modelos facsimile).

Um tema recorrente de discussão é a relação entre os modelos e os dados. Os desafios da verificação e validação (construir o modelo correto), e a sua consistência apoiam-se em testes sólidos usando dados reais. Comparando os resultados e os dados, podemos constatar

que, em muitos casos, as observações podem ser impossíveis e demasiado abstratas. As diferenças podem ser devidas a várias causas, a saber o modelo ser mau, os dados estarem corrompidos, o modelo ser uma abstração do alvo, ou a variações ao acaso, pois as distribuições das amostras podem ser desconhecidas. Hoje em dia, em muitos Congressos, o recurso à apresentação de protocolos ODD (Overview, Design concepts, Details), é obrigatória, pois a especificação precisa dos modelos facilita a replicação dos comportamentos, logo a sua posterior verificação.

A validade envolve dúvidas relacionadas com a sensibilidade, a reputabilidade e a simplicidade. A sensibilidade aos valores dos parâmetros da entrada levanta duas perguntas: Será que as pequenas mudanças nos valores das entradas provocam grandes mudanças nas saídas? Será que é possível ter a certeza que os valores usados são os corretos? A análise da sensibilidade permite correr o modelo várias vezes, com pequenas alterações nos parâmetros da entrada e nas saídas observadas, embora o espaço dos valores da entrada poder ser muito grande. A reputabilidade sugere duas perguntas: É a saída semelhante em cada execução? Se não, será que há a certeza das execuções serem as típicas? Finalmente, a simplicidade também pode levantar duas perguntas: Pode o modelo ser simplificado sem afetarmos a sua validade? Se sim, porquê usar-se um modelo mais complexo do que necessário? Recorre-se amiúde à regra de Occam para se defender sempre as soluções mais simples.

Dois modelos de mentes que vêm sendo usados com frequência são o BDI e o MDP. O BDI (Belief-Desire-Intention) permite um raciocínio qualitativo, apoiado em estados mentais e suporta-se em teorias lógicas. Captura a relação entre componentes de alto nível e segue as estratégias de decisão heurística, baseadas em regras. É bom para grandes problemas (controle de tráfego aéreo). O MDP (Markov Decision Process) permite construir a decisão teórica, dos efeitos para as causas, frequente na tomada de decisão em ambientes competitivos e parcialmente observáveis. Envolve componentes de baixo nível (ações primitivas, recompensas) e políticas associadas a procura de longo prazo, e recorrendo à maximização de uma função de utilidade. É bom para ambientes estocásticos, como as catástrofes, as crises e os mercados.

As nossas mentes organizam, entre outras coisas, a informação envolvida com os atos de comunicação social num formato que pode ser modelado em termos geométricos ou topológicos. Várias têm sido as teorias semânticas sobre o significado das palavras, e uma delas (Gärdenfors 2014) avança os espaços conceituais.

O mapa da cognição da ação (Devlin 1991) esboça como saímos da percepção e chegamos à intenção, através da componente motivacional (dos objetivos aos desejos), da componente informativa (das crenças aos desejos), e da componente deliberativa (das intenções às ações). Embora esta solução seja simples, a sua eficácia justifica-se mesmo em problemas de uma certa complexidade. No entanto, quando visamos problemas sociais e políticos bastam modelos reativos das mente dos agentes, com cognição nula, para obter resultados satisfatórios.

3 Teorias

Existem muitas teorias sobre as mentes (TOM), e sobre a nossa capacidade em atribuir estados mentais a nós próprios e aos outros, e compreender que todos eles podem ser diferentes (Bratman 1987). Estas teorias, por vezes especificações formais, abordam a empatia e a consciência, falam sobre como entendemos e explicamos os comportamentos de terceiros, atribuímos estados mentais, prevemos e deduzimos as ações dos outros, isto é os seus modelos. As teorias são diferentes da Filosofia, pois a mente não é diretamente observável e a mentalidade é um espaço de grande complexidade.

Quando discutimos a mentalidade de um agente artificial somos obrigados a considerar os mecanismos e os dispositivos que temos para construir estas teorias, a estrutura das experiências conscientes, os caminhos que nos levam da agência à intencionalidade, os modos de coordenar e controlar os estados mentais, os modos de perceber as relações entre a percepção, o conhecimento e a realidade, os acessos informacionais dos conhecimentos e das crenças, as mudanças de representação na mente, a distinção entre ver e conhecer, consciência e sono, os estados mentais durante as interações sociais, o registo dos sentimentos e da emoção, em particular, ou ainda os julgamentos sobre os pensamentos e as coisas.

Em 1895, Freud, recorrendo à linguagem, pensava que os processos inconscientes eram mentais, embora os considerasse neurobiológicos. Hoje não estamos tão certos disso, pois os estados podem ser da mente ou do cérebro, e o mental e o biológico coexistem como duais. Mesmo abordando outras formas de comportamento, além da linguagem, a linha de separação é muito ténue. Churchland (2013) acha que a consciência é essencialmente narrativa, e então temos de recorrer à linguagem. Também no esforço em compreendermos os processos do cérebro descobrimos que estes suportam e regulam a consciência. O estar consciente permite a aquisição da linguagem.

O esforço teórico na Inteligência Artificial e na Ciência em geral tem sido notável nos últimos 50 anos. A teoria Sociedade da Mente de Minsky (1988), desenvolvida desde os anos 70 com Seymour Papert, é apenas um primeiro exemplo de reflexão sobre o que é pensar, o que é uma mente, ou seja o que o cérebro faz: “A mente é uma imensa coleção de processos (as entidades que pensam), designados por agentes, que realizam um grande leque de funções, tais como esperar, prever, reparar, lembrar, rever, atuar, depurar, comparar, generalizar, exemplificar por analogia, simplificar, e muitas outras tarefas cognitivas.” A teoria Sociedade da Mente aborda como os grupos se podem organizar em comunidades (via a colaboração) com mais capacidades de que um único agente poderia ter. O autor desenvolveu teorias sobre como os processos (linguagem, memória e aprendizagem) operam, sem esquecer conceitos como a consciência, o sentido do eu, e o livre arbítrio. Minsky, 20 anos depois, refletiu um pouco mais sobre a mente, incorporando os sentimentos em *The Emotion Machine*. Os resultados foram objeto de uma reunião, organizada pela IBM, onde o senso comum esteve no centro de todas as discussões (McCarthy et al 2002).

Os agentes planeiam geralmente num mundo social. Preparam planos de ação, antecipadamente através da comunicação de uns com os outros, e suportam formas complexas de organização para interagir com o ambiente que os envolve e face ao estado futuro do mundo. Os seres humanos, em particular, têm estudado teorias da ação para perceber como isso se desenrola.

A teoria arquitetural CogAff (Sloman 1999) é outro exemplo motivador pois cruza os três níveis de uma mente, ao longo da evolução, a da meta gestão (a mais jovem), referida aos processos reflexivos,

com a do raciocínio deliberativo, referida aos mecanismos “o que se”, e, ainda, com o mais velho dos mecanismos reativos, com as três camadas da percepção, do processamento central, e da ação.

Sabemos que o comportamento de um agente emerge a partir da interação entre dois espaços, o mental e o arquitetural. E por isso as formas das suas arquiteturas podem ser muito diferentes, consoante os estados mentais que escolhemos, os seus atributos, os modos como se compõem (através de leis) e os processos de controle (Corrêa e Coelho 1998).

A maior parte das Teorias da Mente em desenvolvimento começam em questões ingênuas, por exemplo: Como as crianças compreendem as crenças falsas explícitas e as diferenças individuais das preferências? Depois, avançam através de experiências e de hipóteses, que necessitam de confirmação, e levam algum tempo para convencerem, de facto, a comunidade científica. A diversidade das demonstrações, a procura incessante de contra argumentos fortalecem ou destroem para sempre aquilo que parecia ser a verdade. Nas últimas décadas, a simulação tem permitido apoiar algumas destas teorias.

4 Arquiteturas

Quando se faz um levantamento de aplicações, que exigem a presença de agentes artificiais, encontramos uma grande variedade de arquiteturas, desde as reativas, em geral simples, as híbridas, e até as cognitivas, muitas vezes pesadas, multi-camadas e complexas. As arquiteturas são propostas (ou desenhos), organizacionais ou estruturais, sobre a realização (como fazer os modelos) de um agente.

Um agente é uma criatura capaz de perceber o que se passa em redor com os seus sensores, de varrer o ambiente circundante (em busca do que ele necessita de saber), e depois de agir com os seus atuadores, dependendo do poder de controle da sua mente (será a consciência o programa de controle do cérebro?). A sua capacidade de percepção e de cognição (que dispositivos e mecanismos possui) vai ditar depois o seu comportamento, desde os mais triviais, governados por regras (Se-Então) e com ações específicas, até às possibilidades mais exigentes de raciocínio, movimento, planeamento, decisão, comunicação, ou de aquisição, memória e aprendizagem.

No entretenimento, no cinema e nos jogos de computador, podemos recorrer aos atores artificiais (uma espécie de duplos), cada vez mais complexos e com aptidões mais ligadas às tarefas que têm de enfrentar. Eis quatro estruturas possíveis de agentes adequados aos jogos de computador, adotando uma arquitetura em três camadas:

Perceber	Varrer	Gerir (tática, estratégia)	Observar
Raciocinar	Planear	Decidir	Deliberar
Atuar	Comportar	Mover	Executar

A ideia do que é a visão, foi imaginada por David Marr nos anos 70 no MIT quando desenvolvia a sua tese de doutoramento (e, confirmada mais tarde na Neurofisiologia, originando um prémio Nobel), uma cadeia de três níveis de análise (o computacional, o algorítmico, e o implementacional), e essa contribuição foi definitiva para a evolução e, mais tarde, a consolidação da Visão Computacional.

Quando temos várias camadas, cada uma delas encarrega-se de transformar as representações geradas pela camada anterior até que a última produza a resposta. Outras arquiteturas, com mais camadas, como as de Minsky (seis níveis), ou mesmo a de Sloman, explicitam a necessidade das mentes possuírem capacidades mais sofisticadas (reflexão, consciência), além do trivial (Singh et al. 2004) esquema de três partes.

Ora, a resposta de um agente depende da sua sensibilidade para varrer o estado dos acontecimentos (de si próprio, as pós-condições das ações), as ações dos agentes em redor (o conhecimento de si próprio, o que está a fazer, e dos outros, incluindo as mensagens e as observações), e as ações dos objetos inteligentes (importantes para a Internet das coisas). Mas não basta só levar em conta as consequências dos eventos, dos agentes e dos objetos, por vezes é importante atender às respostas emocionais (Trapl, Petta e Payr 2002) e então devemos escolher um modelo de agente, por exemplo o OCC (de Ortony, Clore, Collins), muito adotado na Educação (sistemas tutores inteligentes).

No caso dos agentes reativos, os traços das suas personalidades podem ser descritas por cadeias de bits ou por vetores, e as trocas que efetuam entre si, para satisfazerem as suas necessidades, envol-

vem produtos e recursos. Muitas vezes tentam maximizar uma função de utilidade (lucro, comida, reprodução, conforto, poder) ou de probabilidade, de acordo com uma escolha de preferências: a melhoria da utilidade é alcançada à custa da diminuição de outras utilidades. As interações são realizadas pela adequação dos produtos de cada agente com as necessidades dos outros agentes, e seguindo dinâmicas de troca, de evolução ou de adaptação.

Os modelos dos agentes incluem a caracterização dos seus lados internos e externos, do comportamento (regido por um dispositivo, ou através de uma função que aplica qualquer sequência de percepções numa ação (descrição abstrata) e da história (sequência de percepções). Perante uma certa situação, escolhem-se os modelos dos agentes que possuem propósitos tendo em consideração a mentalidade (escolhas de estados), a racionalidade (maximizar uma função de utilidade), o comportamento (observações, teorias), ou as regras (contexto, fixas, adaptativas, decisão, jogo).

Agente = Estados Internos + Regras Comportamentais

Alguns estados são fixados (sexo, ritmo metabólico, visão) para toda a vida do agente, enquanto outros mudam (preferências económicas, riqueza, identidade cultural, saúde) através da interação entre os outros agentes ou com o ambiente externo, a paisagem (topografia com recursos renováveis) e a rede de comunicações.

Por exemplo, podem-se distinguir (Russell e Norvig 2013) quatro tipos de arquiteturas, em função de um comportamento pretendido: agentes reativos apoiados em regras de produção sem estados internos (sem memória), agentes reativos baseados em modelos com regras de produção com estados internos (com memória), agentes baseados em objetivos com mecanismos de decisão com seleção (de estados) e baseados em objetivos (uso da procura e do planeamento), e agentes baseados em utilidades com mecanismos de decisão racional (funções de utilidade) e habilidade de enfrentar a incerteza.

O modelo mais simples contempla um único ciclo de ação, que começa com as entradas para provocar comportamentos, e com o conhecimento sobre o mundo externo:

Percepção → Controle → Ação

No caso de uma abelha, o que ela pretende é comunicar às outras onde está a comida boa (localização geográfica), logo a mente necessita só de percepção e comunicação. No caso das batérias conversadoras estas pretendem estabelecer alianças para benefício mútuo, e por isso recorrem à linguagem (AI-2) para calcular se o seu número atingiu uma certa densidade (maior que um patamar) e assim decidem se atacam o hospedeiro (sensibilidade ao quórum). No caso de um agente moral, existem relações sociais e regras comportamentais a seguir, e portanto o agente necessita de uma computação capaz de o permitir portar-se bem.

Muitos cientistas têm estudado, ao longo dos anos, a natureza das mentes e a sua organização em camadas ou com módulos. Um deles (Minsky 1995) defendeu que a mente é uma coleção de “kludges” (modos de fazer rápido), aumentada oportunisticamente de outros mecanismos (ativadores, censores, apaziguadores, previsores) à medida, e esta solução parece ser bastante interessante para soluções heurísticas ou de desenrasque.

Nos jogos de computador, os modelos dos agentes adotados são simples, porque temos três necessidades básicas: o movimento, a tomada de decisão sobre os movimentos, e o pensamento tático/estratégico. Uma máquina para jogos, como por exemplo a Eudhora Physics Engine, privilegia o movimento das formas físicas (corpos das personagens), a sua animação e física, apoiado num guião das atividades. Ultimamente, a indústria do entretenimento recorre a agentes mais espertos (com mais neurínios), que necessitam de atuar com independência e autonomia, que se adaptam aos ambientes e exploram convenções, que tomam decisões informadas, que enfrentam a diversidade, que raramente falham e que reagem às mudanças, reorganizando-se, aprendendo por exemplo ou baseando-se na experiência.

O modelo mental BDI (Bratman, Israel e Pollack 1988) inclui três estados mentais (crenças, desejos, intenções), e foi materializado pela arquitetura IRMA para um mundo em grelha, tipo tabuleiro de xadrez, tendo dois ciclos, um de ação e outro de raciocínio. O modelo contempla um comportamento, o do raciocínio prático (dirigido à ação) e dois processos computacionais. No primeiro, decide-se que estado queremos atingir (deliberação) e o resultado é uma escolha.

No segundo, decide-se atingir esse estado (raciocínio meios-fins), e o resultado é um plano. Este mecanismo tem um algoritmo básico (Wooldridge 2002) que é apresentado em seguida:

1. while true
2. observe o mundo;
3. atualize o modelo interno;
4. delibere sobre qual a intenção que deve alcançar a seguir;
5. use o raciocínio meios-fins para obter um plano para essa intenção;
6. execute o plano
7. end-while

O mecanismo separa a atividade de escolher um plano (de uma biblioteca de planos) da execução dos planos que estão correntemente ativos. Por isso, estes agentes são capazes de equilibrar o tempo gasto na deliberação sobre que planos (escolher o que fazer) e executar aqueles planos (fazer). Uma terceira atividade, criar os planos em primeiro lugar (planeamento), não está no âmbito deste modelo, e isso é deixado para o projetista e programador.

O modelo evocado por este algoritmo pode ser traduzido por três componentes simples, articuladas entre si:

Componentes do Ciclo de controle:

Função de revisão de crenças (frc): $f(C) \times Per \rightarrow f(C)$
Processo de deliberação (delibere): $f(C) \rightarrow f(I)$
Raciocínio meios-fins (plano): $f(C) \times f(I) \rightarrow \text{plano}$

Esta arquitetura possui dois níveis do raciocínio prático:

1º Nível: as intenções e os planos anteriores colocam problemas e fornecem um filtro para as opções que são as soluções potenciais desses problemas; e,

2º Nível: as razões desejo-crença entram como considerações para serem depois pesadas (valores) na deliberação entre as opções relevantes e admissíveis.

O modelo BDI apoia-se na Psicologia Popular (“Folk Psychology”) para explicar uma ação. Após aprender os antecedentes causais de certos padrões de comportamento, um agente gera uma nova oportunidade de escolha. Mas, para agir, precisa de vontade, de força. Agir não é só fazer! Muito do que fazemos, e somos, é expresso e explicado em termos do que sentimos, acreditamos, desejamos, tencionamos, esperamos, receamos, etc. A compreensão do comportamento é feita em função de estados mentais, i.e. as atitudes que são as explicações a que uma pessoa recorre para agir. McCarthy disse que “atribuir crenças, livre arbítrio, intenções, consciência, habilidades, ou querer a uma máquina é legítimo quando uma tal atribuição expressa a mesma informação acerca da máquina daquela que é expressa sobre uma pessoa” e, ainda, “é útil quando a atribuição ajuda-nos a compreender a estrutura da máquina, o seu comportamento passado ou futuro, para repará-la ou mesmo melhorá-la.”

Os objetivos são os estados que o agente alcança e as crenças são o que o agente imagina ser o estado do mundo (informação, conhecimento), os desejos, as preferências do agente (motivação), as intenções, os objetivos ou os desejos com os quais o agente se comprometeu trabalhar, as expectativas, as situações em que se espera a ocorrência de algo, e, finalmente, as emoções, ações, modos de pensamento.

Quando se constrói um agente para uma certa paisagem de intervenções, podemos complicar a sua mente, potenciando uma variedade de outros comportamentos (e, de personalidades, traços de cultura e de caráter), não só alargando o seu tipo de pensamento simbólico, mas também incluindo outras atitudes, interesses e sentimentos (emoções). É o que se designa por caracterizar a identidade dinâmica de um agente (Dimas e Prada 2014). Vejamos dois exemplos: os bonecos improvisadores (Improv Puppets, 1987) para o mundo da educação, do grupo da Barbara Hayes Roth da Universidade de Stanford, tinham um corpo e uma mente, e esta possuía um calendário, com controle dos planos, um gestor da agenda dos possíveis (próximos) comportamentos, e um executor; os peixes virtuais 3D, numa piscina virtual da equipe de Dimitri Terzopoulos, da Universidade de Toronto, nos anos 90, tinham percepção (sensores e visão), o que permitia focar as rotinas comportamentais, ativadas também por um gerador de intenções, por sua vez alimentado com os hábitos

particulares daquele grupo de peixes. As rotinas ligavam-se à coordenação motora (atuadores de músculos, formas físicas) e aos tipos de movimentos (ações) pretendidos.

Outros modelos são o BDIE (Corrêa et al 1998), capaz de alargar o número dos estados mentais para quatro atitudes (crenças, desejos, intenções e expectativas), o BVG (Antunes 2001) só com três (crenças, valores e objetivos), organizadas em torno de um cálculo de importância e virado para o lugar das qualidades (recorrendo aos valores, como importância, intensidade, urgência e persistência para fazer as escolhas e tomar decisões ao longo do ciclo de deliberação), o BOID (Belief-Obligation-Intention-Desire) de (Broersen et al 2002), focado no jogo com as normas, ou ainda o BOIDS de Craig Reynolds, onde os comportamentos emergem a partir das interações (conduzidas por regras) entre os vários agentes.

5 Exemplos de instalações e construções sociais

Imaginemos que queremos entender um mercado económico, isto é como várias mercearias, ao longo de uma rua, se comportam (em oligopólio, 2 produtores e n consumidores) face aos seus clientes, ora concorrendo umas contra as outras, ou colaborando entre si, associando-se em cooperativas (tipo Grupa). Podemos estudar a situação fazendo uma experiência com pessoas, separando a procura do fornecimento (produção), antes mesmo de concebermos um aparelho para simularmos, tão perto da realidade, o que se passa. Os dados recolhidos mostram dois estádios, um onde os preços foram baixados por ambos os produtores, e depois um segundo onde os preços tendem a subir até se estabelecer um equilíbrio. Quando começamos a automatizar os processos dos produtores podemos adotar a modelação baseada em agentes (ABM).

Escolheu-se a situação de um mercado oligopolista (Caldas e Coelho 1994), e esse modelo de um agente foi traduzido numa arquitetura de vários módulos, cobrindo a personalidade, a agenda, a assimilação, e a ação. O módulo de ação era um enquadramento (“frame”) composto por número, estratégia preferida, preço, estratégia do competidor, valor da estratégia, percentagem de mercado, tipo, direção de melhoria do lucro, e ação. Na comparação das experiências com sujeitos humanos e com agentes artificiais concluiu-se

que o modelo era demasiado rudimentar (crenças e raciocínio), com 26 regras (tipo ação se condições) e poucas estratégias (competição agressiva, cooperação seguindo um líder ou um seguidor) para estudar os mercados oligopolistas, por causa da ausência de capacidade de aprendizagem (Caldas e Coelho 1994).

A ampliação do modelo anterior para um domínio diferente, o da educação no Ensino Politécnico, permitiu ver até onde podíamos explorar este modelo simples de agente. No caso do ensino da Álgebra por Eusébio (1995) provou-se que necessitávamos de uma melhor capacidade mental, além da simplicidade descrita pelas crenças e por um conjunto de regras de dedução.

Graças ao trabalho de Corrêa (1994), sobre as estruturas de estados mentais, foi possível compreender as potencialidades da mentalidade de um agente (além das crenças, desejos, intenções e expectativas) em situações de conversas entre robôs que queriam passar de uma sala para outra e através de uma porta com fechadura e chave. A via da sociedade de estados começou por ser estruturada numa tábua de estados mentais (cada estado definido por atributos, leis de composição e processos de controle) para os agentes isolados (Corrêa e Coelho 1998), e depois foi alargada para situações coletivas (Corrêa e Coelho 2004).

No trabalho de Eusébio (1995), continuou-se a complicação do modelo do agente, através da introdução dos identificadores dos agentes, dos contextos do domínio do conhecimento (tópicos), das funções parciais de credibilidade de um agentes nos tópicos, das estruturas de dedução (relação parcial de ordem para ter preferências), e das relações entre as fórmulas e os tópicos. Havia uma base de comunicação dos agentes, pois os agentes trocavam mensagens (públicas e privadas), havia assimilação das mensagens, critérios, princípios heurísticos de sinceridade, credulidade e confiança, e deste modo tipos de agentes (arrogantes, envergonhados, vagarosos, sem atenção, e obstinados). Conceberam-se diferentes cenários de aprendizagem, com comunicação orientada para um e dois lados (professor ou alunos). A experimentação permitiu ainda introduzir aspetos inovadores no modelo dos agentes, como o tempo, a intenção, a motivação e a memória.

Na simulação com multidões (formigueiros, rebanhos, cardumes, bandos, enxames, ou “swarms”) de pequenos robôs (Moniz e Urbano

2002), pretendeu-se estudar as estratégias de controle (coordenação) para alcançar consensos em pequenos grupos através da comunicação e da partilha de poder. Tratavam-se de formações de grupos onde se gerava a produção de comportamentos coletivos, globais e complexos, a partir das iniciativas individuais de simples agentes. As transições comportamentais, onde um indivíduo podia liderar ou ser liderado, surgiam também a partir das interações ao nível individual.

No pagamento de impostos (caso da inspeção) existem dois tipos de agentes, contribuintes: amarelos, ou amarelos claros (quando são apanhados a fugir do pagamento de impostos), e inspectores (vermelhos) a tentar apanhá-los. Estes agentes são baseados em utilidades e têm uma regra complexa para decidir quanto podem ou não pagar (evadir): a decisão depende da ponderação da sua riqueza e das multas aplicáveis. Os agentes teimosos são os que decidem sempre da mesma maneira.

No caso de multidões, cardumes, formigueiros, enxames, rebanhos, podem-se descobrir regras simples (algoritmos) que permitem ver as multidões a trabalhar bem. Assim, milhares de animais simples agem como se fossem um cérebro coletivo capaz de tomar decisões e de se mover como um simples organismo. O comportamento pode ser descrito por uma função ou um conjunto de regras.

Podemos escolher novos alvos, como rebanhos de robôs, células no interior do nosso corpo, ou neurónios do nosso cérebro. Os estudos de Iain Couzin (Universidades de Princeton e Oxford) mostraram que os neurónios podem seguir as mesmas regras do comportamento coletivo dos gafanhotos ou mesmo dos peixes. Como é que o cérebro de um indivíduo usa a informação para chegar a uma decisão coletiva sobre o que está a ver? Que regras seguem as células cancerosas para construir tumores ou migrar através dos tecidos?

Podemos filmar um formigueiro ou um enxame de gafanhotos para olhar os comportamentos cooperativos e pacíficos, os de pânico ou de desastre. O comportamento das multidões emerge imprevisivelmente a partir das interações das ações individuais e depende da sua densidade. Como construir modelos de simulação para prever as decisões, e as situações que as enquadram? Como a informação flui através dessa multidão?

Apenas alguns líderes conseguem guiar as multidões de modo efectivo. Dois líderes podem puxar a multidão em direções opostas.

Quando se aumenta a diferença de opinião entre os indivíduos, o grupo espontaneamente chega a um consenso e move-se na direcção escolhida pela maioria.

Podemos adotar agentes pintores para estabelecer a formação descentralizada de consensos (decisão coletiva comum) em sociedades de agentes sem qualquer estrutura organizacional. A cor permitirá visualizar a coordenação de agentes e a seleção de regras sociais que regulam os conflitos inter-individuais (Urbano 2004). Neste caso, um agente será uma máquina de estados de atributos internos, capaz de ter acesso a informação externa (meio ambiente e outros agentes), com a qual alterava os seus atributos.

Num Parque Tecnológico, em Portugal, onde se criaram condições para se estabelecerem ligações entre várias empresas, dispuseram-se no espaço, de modo quase urbanizado, os edifícios de uma universidade, as empresas tecnológicas, as de serviços, e as de equipamentos, os restaurantes, inventando um pouco a ideia de cidade. Tudo isto foi feito sem pensar no lugar mais apropriado para cada uma delas e na ausência de pressupostos de criatividade, inovação, ou mesmo de racionalidade. Os lugares foram atribuídos por razões de preferência do vendedor ou do comprador. Depois, foi-se ver se havia inovação num tal cenário, e descobriu-se que era quase nula.

Na simulação da dinâmica da inovação (apoio à gestão e tomada de decisão), num Parque Tecnológico com 125 empresas (Schilperoord e Coelho 2003), um agente (reativo) era uma estrutura (“frame”, “kene” (knowledge gene), etc.) capaz de incluir os atributos e as propriedades de uma certa entidade. Um “kene” do modelo Intersections era uma cadeia do tipo genético que representava as aptidões e as tecnologias de uma empresa que acumulava ao longo da sua história. As empresas tinham também uma localização geográfica.

O modelo Intersections era um mundo de faz-de-conta sobre os processos de inovação, escrito em Swarm, e destinado a ajudar os decisores a fazerem opções (Schilperoord 2005). Ocorriam dois tipos de interação, 1) estratégias e comportamentos (compatibilidades) para atualizar aptidões e tecnologias, onde o produto era a inovação, e 2) laços sociais (encontros, impactos) da rede onde produto era a sociabilidade. As interações estavam ligadas, local e globalmente na dimensão geográfica (proximidade cultural sobre grafos dinâmicos; e, a inovação como trajetória no espaço, à custa de mecanismos que

faziam sobressair as ligações sociais, os encontros, os acontecimentos, os laços de confiança, os padrões, os agrupamentos e as estratégias (atores, eventos).

Finalmente, no caso dos protestos de rua e nas confrontações violentas, com as forças da ordem, pensou-se uma construção social apoiada num laboratório (ProtestLab) com múltiplos atores, como os manifestantes, os polícias e os jornalistas da media que se podiam encontrar em cenários com pontos de encontro e de atração, obstáculos e entradas/saídas (largo em frente da escadaria de pedra da Assembleia da República em Lisboa). Para simplificar as simulações, e sem perder o realismo, os agentes são reativos e as suas personalidades podem ser descritas por vetores onde os valores podem ser pesados de acordo com a percepção do agente colocado num certo contexto local. Os agentes podem estar em quatro estados diferentes (sossegado, ativo, violento, e combatente). Os tipos de manifestantes podem ser belicoso, disponível e disponível para ver, e com diferentes personalidades. Os polícias podem ter quatro tipos, papel múltiplo, defensivo, ofensivo e comando, também com diferentes personalidades (perfil da missão). Os jornalistas são atraídos pelos episódios de violência, no espaço de protesto, e tentam registar (obter fotos e filmes) desses eventos (Lemos, Coelho e Lopes 2013).

O ProtestLab tem medidas quantitativas dos padrões de multidão, intensidade do protesto, impacto potencial das notícias e desempenho da força de polícia sobre os manifestantes e a construção de um perímetro de defesa. A parametrização e a validação foram alcançadas através da análise da sensibilidade dos atributos dos agentes (raio de visão, pesos do vetor de personalidade), das regras de comportamento (influência dos contextos em mudança) e recorrendo a uma coleção de vídeos de manifestações, aos indicadores dos conflitos sociais e aos inquéritos preenchidos pelos manifestantes em Lisboa (Lemos, Coelho e Lopes 2015).

Os aspetos inovadores do ProtestLab, face a outros trabalhos já publicados, são a representação dos tipos múltiplos dos intervenientes artificiais, com objetivos múltiplos e estados possíveis, e da variedade de interações (efeitos miméticos, movimento evita/aproxima orientado por objetivos, lutas) apenas com um desenho reativo, mas flexível, dos agentes. A arquitetura Varrer-Planear-Comportar ajusta-se bem ao comportamento dos intervenientes escolhidos (manifes-

tantes, polícias, jornalistas), suportado em regras e determinado por objetivos simples, sem necessidade de muito raciocínio e deliberação, embora os alvos e os contextos possam ser múltiplos. Por isto, a parametrização e a validação são mais simples do que em simulações mais sofisticadas com centenas de agentes com aspetos emergentes complexos. Além disso, as dinâmicas realistas foram conseguidas em pequenos computadores, e com a observação fina das situações ocorridas no últimos anos na Europa (Portugal, Espanha, Itália, Grécia, Inglaterra, Ucrânia), no Norte de África (Tunísia, Egipto), na Turquia, e no Brasil.

6 Conjeturas

Chomsky (2012) tentou, numa discussão recente com Norvig, na Revista *The Atlantic*, pôr em causa a importância da Estatística (modelos Bayesianos), face à Lógica de Predicados, na representação do conhecimento (abstração) e na cognição em geral. Os êxitos recentes com o sistema Watson da IBM (e o Siri da Apple) e a sua aplicação futura à Medicina (ligações entre doenças e sintomas: modos de raciocínio, mapas de causalidade), e o lugar e papel desempenhado pelas probabilidades (quantidades numéricas) na ordenação das possibilidades de respostas abrem de novo a discussão sobre os modelos de aprendizagem (racional versus estatística) e a eficácia desta aposta (“Deep Learning”) com milhões de dados.

A IA é neste momento uma anarquia de métodos, de técnicas e tecnologias, um conjunto de vários subcampos quase sempre isolados, enfim uma espécie de Torre de Babel. Serão, hoje em dia, os algoritmos mais importantes do que os modelos e os sistemas de ontem? Será possível integrar as áreas das Redes Neurais com a Computação Evolucionária e a Escola Simbólica da IA Clássica? E, o que fazer das Neurociências e da Ciência Cognitiva, se quisermos explorar também os aspetos mais pertinentes da escola simbólica, a GOFAI (Good Old Fashion AI)?

O exemplo do senso comum pode levar à especulação feita por Yann LeCun (Facebook) de que a aprendizagem não supervisionada e algum treino pode conduzir a previsões sobre o que vem a seguir, se forem conhecidas de antemão as restrições do mundo físico. O uso de redes de memória, de redes neuronais e de aprendizagem profun-

da tem permitido alguns avanços na procura com grandes quantidades de dados (em textos, imagens e vídeos).

Churchland (2013) perguntou “O que é que o cérebro está a fazer quando nós estamos totalmente conscientes, e quando não o estamos? O que se passa entre o sonhar, o estar em sono profundo e o estar consciente?” E, depois, comparou os avanços dos neurocientistas nos últimos anos em várias frentes com os pensamentos dos filósofos, para concluir que os segundos achavam que esses avanços seriam impossíveis. Ou seja, mesmo compreendendo tão pouco sobre o cérebro (Qual é o lugar da consciência no cérebro? E, o da memória, do auto-controle, ou ainda o da organização da ação? O que guia a atenção do topo para a base? Como é que o cérebro sabe o que ignorar e o que focar?) a ciência continua afastando-se do senso comum e a avançar em passos seguros.

7 Conclusões

A ideia de mente está associada a algo que é capaz de perceber e pensar, ou seja compreender e raciocinar. Não está isolada, mas vigilante e ligada ao cérebro e ao corpo, acompanhando os comportamentos, tratando dos sentimentos e fazendo experiências para conhecer sempre mais e conseguir ainda equilibrar o objetivo e o subjetivo.

Muitas vezes, o trabalho envolvido pela concepção de uma mente começa com uma ideia ingénua e simples. Depois, vira numa espécie de jogo, uma interação entre interiorizações intuitivas, modelação teórica, realizações práticas, estudos empíricos, experimentações, e análises científicas. A intuição está muito perto do pensamento criativo, a modelação recorre à matemática, a construção é engenharia e tentativa-e-erro (“hacking”), enquanto o estudo empírico e a análise são a ciência. Este vai e vem faz que as idealizações da mente abram espaços significativos, para os avanços e recuos de hipóteses, as discussões acaloradas com troca de argumentos, e mais tarde para a instalação de aparelhos experimentais destinados a observar até que ponto o esforço visionário estava certo.¹

¹ Agradecimentos: Este trabalho foi apoiado por uma verba (para o BioISI, referência da Unidade: UID/MULTI/04046/2013), da FCT/MCTES/PIDDAC, Portugal.

Helder Coelho
BioISI e FCUL

Referências

- Anderson, J. R. 2014. *Cognitive Psychology and Its Implications*. Worth Publications, 8th Edition.
- Antunes, L. 2001. *Agentes com Decisões Baseadas em Valores*. Tese de Doutoramento (Orientador: H. Coelho), FCUL, Lisboa.
- Antunes, L., Nunes, D. e Coelho, H. 2014. The Geometry of Desire. *Proceedings of AAMAS-14*, Paris, May 9.
- Baum, E. B. 2004. *What is Thought?* Cambridge, MA: The MIT Press.
- Boden, M. 2004. *The Creative Mind: Myths and Mechanisms*. Routledge, Taylor and Francis Group, 2nd Edition.
- Bratman, M. E. 1987. *Intentions, Plans, and practical Reason*. Harvard University Press.
- Bratman, M. E., Israel, D. e Pollack, M. 1988. Plans and Resource-Bounded Practical Reasoning. *Computational Intelligence* 4(3): 349-355.
- Bratman, M. E. 1999. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press.
- Broersen, J., Dastani, M., Hulstijn, J. e van der Torre, L. 2002. Goal Generation in the BOLD Architecture. *Cognitive Science Quarterly* 2(3-4): 428-447.
- Caldas, J. M. C. e Coelho, H. 1992. Strategic Interaction in Oligopolistic Markets – Experimenting with Real and Artificial Agents. *Proceedings of the Fourth European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, Rome July 29-31. (e *Artificial Social Systems*, C. Castelfranchi and E. Werner (Eds.), Lecture Notes in AI 830, Springer-Verlag, 1994).
- Caldas, J. M. C. e Coelho, H. 1998. Teamwork by Organised Populations: Private Benefits, Public Costs. *Proceedings of the II Conference Complexity in Economics, Cycles and Chaos in Economics*, ISEG, May 15-16.
- Caldas, J. M. C. 2000. Escolha e Instituições: Uma Análise Económica Aplicada na Simulação Multiagentes. Tese de Doutoramento, ISCTE (Orientador: H. Coelho), Lisboa.
- Churchland, P. 1989. *Neurophilosophy: Toward a Unified Science of the Mind/Brain* (Computational Models of Cognition and Perception). Cambridge, MA: MIT Press.
- Churchland, P. 2013. *Touching a Nerve, Our Brains, Our Selves*. New York: W. W. Norton & Company.
- Corrêa, M. 1994. *A Arquitetura de Diálogos entre Agentes Cognitivos Distribuídos*. Tese de Doutoramento, UFRJ (Orientador: H. Coelho), Rio de Janeiro.
- Corrêa, M. e Coelho, H. 1998. From Mental States and Architectures to Agent's Programming. *Proceedings of the Iberoamerican Congress on Artificial Intelligence (IBERAMIA'98)*, Lisbon, October 6-9. (Springer-Verlag Lecture Notes in AI 1484, pp. 64-75, 1998).
- Corrêa, M. e Coelho, H. 2004. Collective Mental States in an Extended Mental States Framework. *Proceedings of the Conference on Collective Intentionality II*, Corsa di Pontignano (Italy), October 13-15.
- Corrêa, M. e Coelho, H. 2010. Abstract Mental Descriptions for Agent Design. *Intelligent Decision Technologies* 4(2): 115-132.
- De Bono, E. 1976. *Mechanisms of Mind*. London: Penguin Books.

- De Waal, F. 2006. *Our Inner Ape: The Best and Worst of Human Nature*. London: Granta Books.
- Decety, J. e Weatley, T. 2015. *The Moral Brain: A Multidisciplinary Perspective*. Cambridge, MA: MIT Press.
- Devlin, K. 1991. *Logic and Information*. Cambridge University Press.
- Dimas, S. e Prada, R. 2014. Dynamic Identity Model for Agents. In S. J. Alam e H. Van Dyke Parunak (Eds.), *MABS2013, LNAI 8235*. Springer-Verlag.
- Edmonds, B. e Meyer, R. 2013. *Simulating Social Complexity. A Handbook*. Springer.
- Epstein, J. M. 2007. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press.
- Epstein, J. M. e Axtell, R. L. 1996. *Growing Artificial Societies: Social Science from the Bottom Up* (Complex Adaptive Systems). Brookings Institution Press.
- Eusébio, A., Costa, E. e Coelho H. 1995. A Closer Look to Artificial Learning Environments. *Proceedings of the 7th EPIA*, Funchal, 3-6 October, Springer-Verlag Lectures Notes on AI 990.
- Fodor, J. A. 1975. *The Language of Thought*. Cambridge, Mass.: Harvard University Press.
- Fodor, J. A. 1980. Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology. *Behavioral and Brain Science* 3(1): 63.
- Fodor, J. A. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, Mass.: MIT Press.
- Fodor, J. A. 1993. Unpacking a Dog. Review of Christopher Peacocke's *A study of concepts*. *London Review of Books* 15 (19): 14-15.
- Freud, S. 1895. Projeto para uma Psicologia Científica. In *Obras Psicológicas Completas De Sigmund Freud*, Vol. I, Rio de Janeiro, Imago Editora, 3ª Edição, 1985.
- Gärdenfors, P. 2014. *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. Cambridge, Mass.: MIT Press.
- Gilbert, N. 2007. *Agent-Based Models* (Quantitative Applications in the Social Sciences). London: Sage Publications.
- Gilbert, N. e Troitzsch, K. G. 2005. *Simulation for the Social Scientist*. Open University Press, 2nd Edition.
- Girard, R. 1965. *Deceit, Desire, and the Novel: Self and the Other in Literary Structure*. The John Hopkins University Press.
- Haugeland, J. (Ed.) 1988. *Mind Design, Philosophy, Psychology, Artificial Intelligence*. The MIT Press, 6th Edition.
- Hawkins, J. 2005. *On Intelligence*. St. Martin's Press.
- Helbing, D. (Ed.) 2012. *Social Self-Organization, Agent-Based Simulations and Experiments to Study Emergent Social Behavior*. Springer.
- Kurzweil, R. 2012. *How to Create a Mind: The Secret of Human Thought Revealed*. Viking Press.
- Lemos, C., Coelho, H. e Lopes, R. J. 2013. Agent-based Modeling of Social Conflict, Civil Violence and Revolution: State-of-the-art-review and further Comments. *Proceedings of EUMAS 2013*, Toulouse, 12-13 December.
- Lemos, C., Lopes, R. J. e Coelho, H. 2014. Quantitative measures of emergent properties in Agent-Based models of street protests and violent confrontation. *Proceedings of the Brazilian Workshop on Social Simulation (BWSS-2014)*, November 4, in *5th World Congress on Social Simulation*. São Paulo (Brazil), IEEE Press, November 4-7.
- Lemos, C., Coelho, H. e Lopes, R. J. 2014. Agent-Based Modeling of Protests and Violent Confrontation: A Macro-Situational, Multi-Player, Contextual Rule-based Approach. *5th World Congress on Social Simulation*. São Paulo (Brazil), November 4-7 November.

- Lemos, C. Coelho, H. e Lopes, R. J. 2014. An Agent-Based Model of Civil Violence with Imprisonment Delay and Legitimacy Feedback. *2nd World Conference on Complex Systems (WCCS14)*, Agadir (Marrocos), November 10-12.
- Lemos, C., Coelho, H. e Lopes, R. J. 2015. *ProtestLab – An Agent-Based Tool for Simulation of Street Protests and Violent Confrontation* (submitted to JAAMAS).
- Lemos, C., Lopes, R. J. e Coelho, H. 2015. *On Legitimacy Feedback Mechanisms in Agent-Based Modelling of Civil Violence* (submitted to International Journal of Intelligent Systems).
- Lenat, D. 1995. CYC: A Large-scale Investment in Knowledge Infrastructure. *Communications of the ACM*, Volume 38, Issue 11, November.
- Markram, H. 2006. Blue Brain Project. *Nature Reviews Neuroscience* 7: 153-160
- McCarthy et al. 2002. An Architecture of Diversity for Commonsense Reasoning. *IBM Systems Journal* 41(3).
- Minsky, M. 1988. *The Society of Mind*. Simon and Schuster.
- Minsky, M. 1995. A Framework for Representing Knowledge. In *Computation and Intelligence*. Luger, G. F. Edition, AAAI Press.
- Minsky, M. 2006. *The Emotion Machine, Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. Simon & Schuster.
- Moniz, L. e Urbano, P. 2002. *Experiências com Agentes*. Working Report, FCUL.
- Newell, A. 1994. *Unified Theories of Cognition*. Harvard University Press, 2nd Edition.
- Nilsson, N. 2009. *The Quest for Artificial Intelligence*. Cambridge University Press.
- Putnam, H. 1961. Brains and Behaviour. *American Association for the Advancement of Science*. Section L (History and Philosophy of Science), December 27.
- Russell, S. e Norvig, P. 2013. *Artificial Intelligence: A Modern Approach*. 4th Edition.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson's University Library.
- Schilperoord, M., Caldas, J. M. C. e Coelho, H. 2003. INTERSECTIONS: A Management Tool for Science and Technology Parks. *Proceedings of the International Conference on Science and Technological Parks*, Lisbon, June 1-4.
- Schilperoord, M. 2005. *Complexity in Foresight, Experiences with INTERSECTIONS: An Agent-Based Simulation Workbench to Help Achieve Adaptiveness in Strategic Planning*. Erasmus University (Co-Orientador: H. Coelho), Rotterdam.
- Singh, P., Minsky, M. e Eslik, I. 2004. Computing Common Sense. *BT Technology Journal* 22(4), October.
- Sloman, A. 1999. What Sort of Architecture is Required for a Human-Like Agent. In *Foundations of Rational Agency*. Wooldridge, M. and Rao, A. (Eds.), Applied Logic Series (14), Springer.
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L. e Goodman, N. D. 2011. How to Grow a Mind: Statistics, Structure, and Abstraction. *Science* (331), March 11.
- Trappl, R., Petta, P. e Payr, S. 2002. *Emotions in Humans and Artifacts*. The MIT Press.
- Urbano, P. 2004. *Jogos Descentralizados de Consenso, ou de Consenso em Consenso*. FCUL (Orientador: H. Coelho), Lisboa.
- Vendler, Z. 1984. *The Matter of Minds*. Oxford University Press.
- Wooldridge, M. 2002. *An Introduction to MultiAgent Systems*. John Wiley.